

Durham Research Online

Deposited in DRO:

26 April 2020

Version of attached file:

Accepted Version

Peer-review status of attached file:

Peer-reviewed

Citation for published item:

Wang, Q. and Megherbi, N. and Breckon, T.P. (2020) 'A reference architecture for plausible Threat Image Projection (TIP) within 3D X-ray computed tomography volumes.', *Journal of X-ray science and technology*, 28 (3). pp. 507-526.

Further information on publisher's website:

<https://doi.org/10.3233/XST-200654>

Publisher's copyright statement:

The final publication is available at IOS Press through <https://doi.org/10.3233/XST-200654>

Additional information:

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full DRO policy](#) for further details.

A Reference Architecture for Plausible Threat Image Projection (TIP) Within 3D X-ray Computed Tomography Volumes

Qian Wang *Member, IEEE*, Najla Megherbi, Toby P. Breckon *Member, IEEE*

Abstract—Threat Image Projection (TIP) is a technique used in X-ray security baggage screening systems that superimposes a threat object signature onto a benign X-ray baggage image in a plausible and realistic manner. It has been shown to be highly effective in evaluating the ongoing performance of human operators, improving their vigilance and performance on threat detection. However, with the increasing use of 3D Computed Tomography (CT) in aviation security for both hold and cabin baggage screening a significant challenge arises in extending TIP to 3D CT volumes due to the difficulty in 3D CT volume segmentation and the proper insertion location determination. In this paper, we present an approach for 3D TIP in CT volumes targeting realistic and plausible threat object insertion within 3D CT baggage images. The proposed approach consists of dual threat (source) and baggage (target) volume segmentation, particle swarm optimisation based insertion determination and metal artefact generation. In addition, we propose a TIP quality score metric to evaluate the quality of generated TIP volumes. Qualitative evaluations on real 3D CT baggage imagery show that our approach is able to generate realistic and plausible TIP which are indiscernible from real CT volumes and the TIP quality scores are consistent with human evaluations.

Index Terms—threat image projection, X-ray computed tomography, CT volume segmentation, baggage security screening, particle swarm optimisation, TIP quality score

I. INTRODUCTION

Scanning passenger baggage using X-ray technology is a mandatory process in airports and other public transportation for security. Although automatic threat material and prohibited item detection using advanced machine learning techniques have been studied [1], [2], [3], they have not yet achieved maturity whereby human operators can be completely replaced. As such the performance of human operators can vary depending on experience, fatigue and baggage item complexity. Threat Image Projection (TIP) is a technique applied in X-ray image based baggage screening systems to monitor the ongoing performance of human operators. TIP is used to generate plausible and realistic X-ray baggage images containing threat signatures (e.g., firearms, improvised explosive devices, etc.) by projecting fictional threat object images onto X-ray images of real passenger bags present within the live aviation security process. It is a little known

fact that TIP is a legally mandated process by both national and international aviation security regulations [4], [5].

Using TIP in X-ray security scanners has been shown to be effective in improving the vigilance and attention of human operators, hence improving the overall performance of threat detection [6], [7]. The benefits of TIP systems are multi-fold. TIP systems make it possible for operators to encounter baggage images with threat objects more frequently during their regular working patterns by randomly applying TIP to benign passenger bags so that they can get familiar with potential real yet rare threats in order to improve their detection ability [7]. Research also suggests operators are motivated and more attentive to do well when knowing that TIP systems are enabled and their performance is monitored [4]. In addition, TIP systems record the performance of individual operators such that this information could be further analysed and used to customize training plans.

These benefits, however, are only achievable if TIP systems are properly used and managed [7]. For example, it is important how frequently an operator should be exposed to TIP during their work pattern. On the other hand, the management of TIP library is critical to the effectiveness of TIP systems. According to [4], the TIP library shall contain a minimum of 1,000 virtual images and 250 threat objects captured in different orientations. The TIP library needs to be updated each year with no fewer than 100 virtual images replaced with new ones. To satisfy these requirements, an effective algorithm of plausible and realistic TIP image generation is crucial.

Recently, 3D Computed Tomography (CT) scanners have seen increasing deployment in airports for baggage screening [8]. A recent study in [9] shows the superior performance of threat detection using 3D CT imaging against traditional 2D X-ray images. However, 3D TIP within CT volumes is still a very challenging problem due to a number of additional factors. Firstly, some form of 3D CT volume segmentation is required to both isolate the bounds of the 3D threat object (source) and the exterior boundary and internal void regions of the 3D baggage item (target). Most CT volume segmentation algorithms are targeted at medical images [10] which can not be readily applied to baggage volumes [11]. Secondly, inserted threat objects have to avoid intersection with existing items in the target baggage volume and additionally exhibit artefacts consistency with those of the original scanned objects already present therein. [12].

To address aforementioned issues in 3D TIP, we extend the work in [12] and present a novel approach for fully automatic

Qian Wang is with Department of Computer Science, Durham University, UK. E-mail: qian.wang173@hotmail.com;

Najla Megherbi is with School of Engineering, Cranfield University, UK. E-mail: najla.megherbi@gmail.com;

Toby P. Breckon is with Department of Engineering and Department of Computer Science, Durham University, UK. E-mail: toby.breckon@durham.ac.uk.

threat image projection within 3D CT security imagery. Our approach consists of four components: threat isolation, void determination, object insertion optimisation and metal artefact generation. Specifically, a threat volume is segmented into the background, threat body and uncertain regions (threat isolation), whilst a baggage volume is segmented into background, inner-void and bag-content regions (void determination). The segmentation results are used to evaluate the quality of a given insertion location and orientation. The optimal insertion is derived by particle swarm optimisation (object insertion optimisation). Finally, metal artefact generation [13] is applied to the generated TIP to enhance the plausibility. In summary, the paper has the following contributions:

- it is among the first attempts to address the threat image projection in 3D CT volumes to our best knowledge;
- the proposed framework integrates 3D object segmentation and particle swarm optimisation algorithms towards the generation of realistic and plausible threat image projection;
- the proposed approach has been validated on real baggage data collected from airports and the experimental results demonstrate its effectiveness from both qualitative and quantitative perspectives.

The remainder of this paper is structured as follows: prior works relevant to 2D and 3D TIP are reviewed in Section II; we present details of our approach for 3D TIP in Section III; qualitative evaluations of each component in our approach are given in Section IV; finally, we discuss limitations existing in the current approach and potential directions of future work in Section V and conclude the paper in Section VI.

II. RELATED WORK

In this section, we make a thorough review of TIP related works including those focusing on 2D and 3D imagery.

A. 2D TIP

The concept of threat image projection in X-ray imagery dates back to 1990s towards the enhancement of X-ray baggage screening performance in airports [14], [15]. As TIP within 2D imagery is essentially to superimpose a threat X-ray image onto a baggage X-ray image in a random position, it is a relatively simple technique in terms of contemporary image processing. As a result, literature on 2D TIP mainly focused on the system design [16], [17] or performance evaluation [6], [7], [14], [15] rather than image processing details. For example, Neiderman et al. [16] designed and patented a means for training and testing baggage screening operators using the 2D TIP technique. An exception is [17] in which the authors presented the details of combining distorted threat images with baggage images to generate realistic and diverse TIP. In addition, our recent work [18], focusing on the investigation of TIP based data augmentation for object detection in X-ray baggage images, also provides some details of viable 2D TIP gleaned from various obfuscated sources [16], [17], [19].

Except for baggage screening, 2D TIP was also applied in video surveillance [20], [21] and cargo screening systems [19]. Neil et al. [20] discussed the challenge and plausibility

of applying TIP in video surveillance. Donald et al. [22] further discussed how TIP (or IGO, Inserted Graphic Objects) could improve vigilance performance, target incident detection rate, and design considerations for TIP images in video surveillance. Quantitative evaluations conducted by Donald [21], however, disclosed IGO were not effective in enhancing the detection of significant events in video surveillance. The reason, as discussed by the author, could be multifold and applying TIP in video streams is quite different from doing it in X-ray images.

Rogers et al. [19] applied X-ray TIP techniques in cargo screening tasks. They proposed a framework extracting threat masks from X-ray images and projecting them onto benign X-ray images to generate realistic TIP. Quantitative evaluations indicated the generated TIP and real X-ray images containing threats were indistinguishable. In addition, transformations were made to inject variation into the threat signatures to generate a very large number of realistic TIP data for training deep learning based object detection algorithms. Among seven types of transformations employed in [19], threat insertion position and rotation are also used in our approach. The employment of such transformation in our approach aims to improve the plausibility of generated TIP which is not a problem in 2D TIP, while in [19], the transformation aims to diversify generated TIP for data augmentation.

B. 3D TIP

Early attempts were made to extend 2D X-ray TIP to 3D TIP but unsuccessful due to obvious visual imaging artifacts which could provide cues for scanner operators to readily recognize the presence of TIP [23]. To address this issue, Yildiz et al. [23] proposed to project threat objects into the sinogram space instead of the original imagery space. They claimed plausible TIP could be generated but their insertion positions were manually decided and the evaluations were conducted on uncluttered image examples without explicit metal artefact generation. Megherbi et al. [12] proposed an approach to fully automatic 3D TIP and applied it to densely cluttered 3D CT baggage volumes (which is subsequently identified as prior work in the commercial implementations of [24], [25]). The approach consisted of three main components: void determination, object insertion location determination and metal artefacts generation. The work presented in this paper is based on [12] but with notable variations and improvement within the stages of threat isolation and object insertion.

III. AUTOMATIC 3D THREAT IMAGE PROJECTION

Our approach to 3D threat image projection is composed of four parts: threat isolation, void determination, object insertion optimisation and metal artefact generation. The framework of our approach is illustrated in Figure 1.

Threat isolation aims to segment threat objects from the background in threat volumes \mathcal{I}^{thr} . These threat objects of interest are prepared beforehand and scanned in a controlled condition (e.g. background voxels with lower values than threat object voxels) for easy segmentation. The subsequent

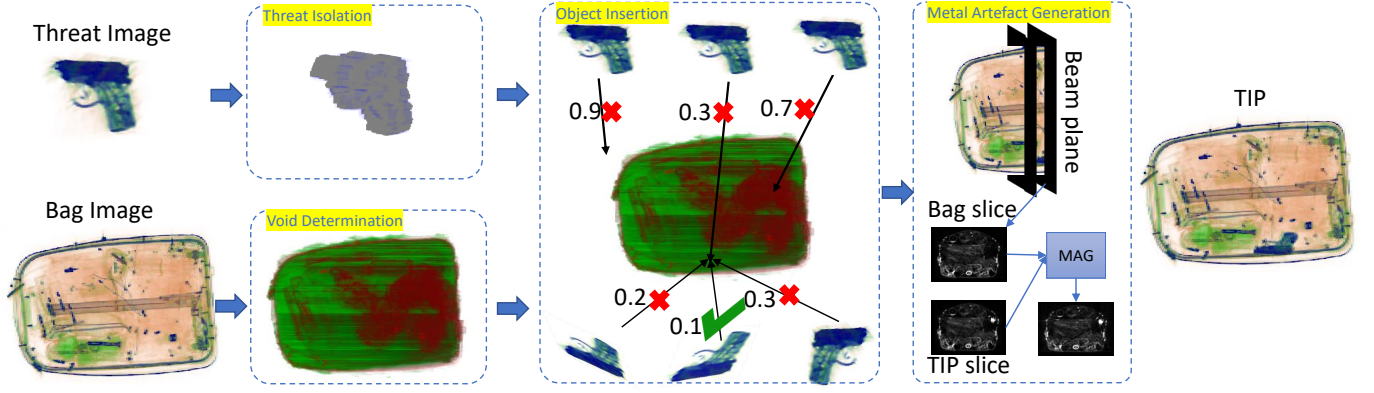


Fig. 1. The framework of proposed 3D threat image projection approach; given a threat CT volume and a baggage CT volume as inputs, a plausible TIP is generated as the output of the approach with the pipeline consisting of four components: threat isolation, void determination, object insertion optimisation and metal artefact generation.

thresholding and morphological operations used in our approach will be described in the following subsection.

Void determination aims to segment baggage CT volumes \mathcal{I}^{bag} into three regions: outer region, inner void and bag content. Different costs will be incurred when the threat object is projected into these three regions. As a result, bag volume segmentation results in a projection cost map of the bag volume indicating the cost of voxels onto which the threat object is projected. With the segmented threat object and the projection cost map of a bag volume, the insertion is boiled down to an optimisation problem which aims to find optimal insertion locations and threat object orientations. Particle swarm optimisation [26] is used in our approach as one of the enabling techniques. To enhance the plausibility of the generated TIP volume, we apply metal artefact generation [13] as post processing to generate plausible metal artefacts in the TIP volume.

In the following subsections, we will present four parts of our threat image projection approach in more detail.

A. Threat Isolation

A variety of threat objects including firearms and improvised explosive devices could be used to generate TIP. To make segmentation easy and accurate, we assume the threat objects are scanned in controlled conditions. Every threat object will be scanned individually with only low-density supporting objects (e.g., foam) if necessary. As a result, threat object volumes (source) are almost free of noise except when metal components exist in the threat objects themselves. Special care needs to be taken to get rid of any residual artefact noise surrounding the body of the threat object.

The pipeline of our threat isolation is shown in Figure 2. It takes a threat volume as the input and outputs a cropped 3D volume of the threat object which is ready to be inserted in a benign bag volume target for TIP generation. CT volumes contain noise with small non-zero voxel values in the void region. To remove the effects of such noise, in the first step of threat isolation, we set a threshold value to binarise the input CT volume. Whilst this simple thresholding process will isolate the threat object in most cases, there could be special

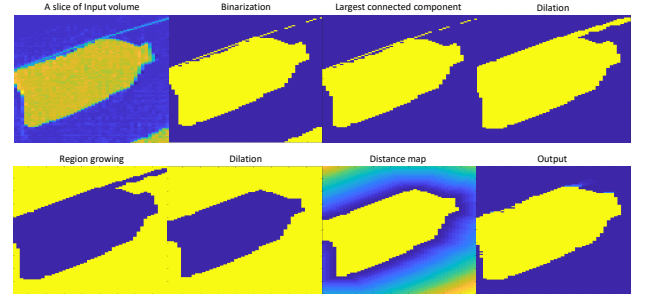


Fig. 2. The pipeline of threat isolation algorithm; slices of a CT volume of a source item bottle is used here for illustration while the segmentation algorithm is actually applied to 3D CT volumes which are segmented into three regions: the threat body region in yellow, the background in blue and the uncertain region in a gradient colour.

cases where the threat objects have an internal sub-void which should be considered as a part of the threat object. In addition, there could exist more significant noise in the CT volume which can not be removed by such simple thresholding. To handle these special cases, we develop a robust threat isolation approach advancing upon that of [12].

Specifically, connected component labelling (CCL) [27] is applied to the binary volume derived from the thresholding process. The resulting labelled connected components could belong to either the threat object or background noise. Due to the fact that the noise components have far fewer voxels than the threat object, we only reserve the largest labelled connected component as the segmented threat object. As such we have successfully removed the noise in the volume but the resulted threat object can still have an internal void. To ensure that internal voids are treated as a part of the threat object, we try to conversely determine the exterior boundary of the threat object. Subsequently, the threat object including the possible void space inside can be derived accordingly. For this purpose, we use a region growing method [28] to segment the exterior region in a threat volume. The region growing seed is usually set to the upper-left upper-leftmost voxel such that this will not be a threat object voxel within a controlled condition CT image scan. To use region growing the threat boundary should

be closed so that the region cannot mistakenly grow into the threat object. To this end, we apply morphological dilation to the isolated threat component derived from the connected component labelling in the previous step.

Region growing is able to segment the non-threat (background) region in the threat volume (source). To reduce the noise surrounding the threat body, a dilation operation is applied to the non-threat region. The dilation operation transforms the voxels close to the threat boundary into the background and effectively removes some noisy voxels from the threat object. However, it could also lead to damage to the true voxels of the threat object. To alleviate this issue, we consider the voxels removed by the background dilation form an uncertain region since the voxels of this region could belong to the threat object or be noise.

Now the threat volume is segmented into three parts: the threat object, the uncertain region and the background. A minimum 3D volume of the threat object is cropped and most of the background region is removed. To facilitate the presentation, we use \mathbf{I}^{thr} to denote the cropped 3D volume of a threat object and a 3D indicator matrix \mathbf{M}^{thr} is used to represent the segmentation results. The element value of $\mathbf{M}^{thr}(i, j, k)$ is determined as follows:

$$\mathbf{M}^{thr}(i, j, k) = \begin{cases} 1, & \text{threat object,} \\ 1/d_{ijk}^2, & \text{uncertain region,} \\ 0, & \text{background.} \end{cases} \quad (1)$$

where d_{ijk} is the distance of voxel (i, j, k) to the boundary of the threat object, which can be calculated by distance transform method proposed by Maurer et al. [29]. Equation (1) results in a 3D volume composed of three different regions: threat body voxels indicated by ones, background voxels with zero values and uncertain voxels in the range of $0 - 1$. The indicator matrix \mathbf{M}^{thr} will serve as a weight matrix to extract the threat object from the original CT volume for TIP generation. The uncertain voxels far from the threat object will have lower weights so that the sharp transition effect can be alleviated when inserting the threat into a benign bag volume. As a result, the resultant TIP look more plausible by using the indicator matrix in Eq. (1) for initial threat isolation.

B. Void Determination

To insert the segmented threat into a plausible location in the bag volume, we require to understand different regions in the bag volume. We propose a bag volume segmentation method to segment a bag volume into: *outer-bag* (background), *bag-content* and *inner-void* regions. Similar to the threat volume segmentation, the pipeline of bag volume segmentation is composed of several morphological operations as illustrated in Figure 3. The pipeline takes a 3D CT volume as input and outputs a indicator matrix representing the segmentation results.

A similar approach to threat volume segmentation is used here to segment the background region of bag volumes. The original CT volume is firstly binarised by thresholding. Subsequently, the largest connected component is extracted by 3D CCL as the volume of a bag which is further dilated to

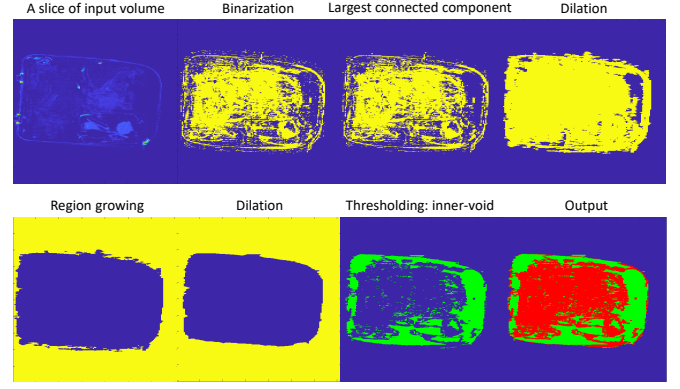


Fig. 3. The pipeline of void determination algorithm (left to right, top to bottom); slices of a 3D CT volume of a suitcase are used for illustration while the algorithm is actually applied to 3D CT volumes which are segmented into three regions: the outer-bag region in blue, the inner-void region in green and the bag-content region in red.

ensure the boundary is closed. The background is segmented by region growing from a random seed (usually set as the upper left most voxel) outside the volume of the bag.

To segment a volume of bag into an *inner-void* region and a *bag-content* region, a simple thresholding is applied so that voxels of smaller values than the threshold form the *inner-void* region and others form the *bag-content* region.

We use \mathbf{I}^{bag} to denote the 3D volume matrix of a bag volume \mathcal{I}^{bag} and subsequently the segmentation results can be represented by a 3D projection cost map \mathbf{M}^{bag} as follows:

$$\mathbf{M}^{bag}(i, j, k) = \begin{cases} 0, & \text{inner-void,} \\ \tilde{v}_{ijk}, & \text{bag-content,} \\ c, & \text{outer-bag.} \end{cases} \quad (2)$$

where c is a large positive constant indicating that a big cost will be incurred if threat object is projected to the *outer-bag* region, and $\tilde{v}_{ijk} = \mathbf{I}^{bag}(i, j, k) \times c/m$ denotes the cost of projecting onto *bag-content* voxels which is equal to the normalized intensity value of voxel (i, j, k) where m is the maximum voxel density value in the CT volume.

C. Object Insertion Optimization

Inserting a segmented threat object into a benign bag in the 3D CT imagery involves the determination of optimal insertion locations and orientations. To enable plausible and realistic TIP, it is important to find suitable locations in a benign bag and proper orientations of the threat object. Ideally, we tend to insert a threat object into the *inner-void* region in the bag volume. In cases where a large void region is available in the benign bag, we need to consider the effect of gravity and insert it into a lower position so that the inserted threat object will not appear to implausibly levitate unsupported. In practice, however, most baggage is cluttered without enough void space for big threat objects. In these cases we allow the threat object to be inserted into regions where voxel intensities are low. Regions of low voxel values are usually occupied by clothes and it is highly plausible to have a threat object concealed in clothes within baggage. In our TIP framework, we formulate

an optimization problem to ensure optimal insertions and use Particle Swarm Optimization (PSO) [26] to solve the problem. Once the optimal location and orientation have been derived, a simple blending approach is employed to project the threat object into the bag volume and generate a final TIP volume output.

1) *Optimizing Insertion Location and Orientation:* Finding an optimal insertion location and orientation is essentially an optimization problem. The variables to optimize are denoted by $\Theta = \{x, y, z, \alpha, \beta, \gamma\}$, where x, y, z are the insertion coordinates within the volume of bag (target) and α, β, γ are the rotation angles of the threat object (source) within three planes (yz, xz and xy). We use \tilde{M}^{thr} to denote the 3D volume matrix after rotating M^{thr} by angles of α, β, γ within three planes.

$$\tilde{M}^{thr} = f(M^{thr} | \alpha, \beta, \gamma) \quad (3)$$

where $f(\cdot)$ is a rotation function which could be implemented by spline interpolation [30], [31]. We use \tilde{M}^{bag} to denote a 3D volume cropped from the projection cost map M^{bag} of the bag volume. The cropping is conditioned on the coordinates x, y, z and the size of rotated threat volume \tilde{M}^{thr} :

$$\tilde{M}^{bag} = g(M^{bag} | x, y, z, \Delta_x, \Delta_y, \Delta_z) \quad (4)$$

where $g(\cdot)$ represents the cropping process, $\Delta_x, \Delta_y, \Delta_z$ are sizes of the cropped 3D volume and equal to those of the rotated threat volume \tilde{M}^{thr} .

The objective of the optimization problem is formulated as follows:

$$\min_{\Theta} \|M^{tip}\|_1 + \lambda_1 \|\tilde{M}^{tip}\|_1 + \lambda_2 y \quad (5)$$

where

$$\tilde{M}^{tip}(i, j, k) = H(M^{tip}(i, j, k) - c') \quad (6)$$

and

$$M^{tip} = \tilde{M}^{thr} \odot \tilde{M}^{bag}. \quad (7)$$

$H(\cdot)$ in Eq. (6) is a unit step function applied to all elements of M^{tip} so that $\tilde{M}^{tip}(i, j, k) = 1$ when $M^{tip}(i, j, k)$ is greater than the constant parameter c' , otherwise 0. The operator \odot between two matrices in Eq. (7) denotes the Hadamard product. The operator $\|\cdot\|_1$ in Eq. (5) is the entrywise matrix 1-norm which calculates the sum over the absolute values of all elements in a matrix; λ_1 and λ_2 are two hyper-parameters adjusting the weights of different terms in the objective function.

Minimizing the first term in Eq. (5) ensures the threat object to be inserted in a region with the lowest average voxel intensity. However, it could result in a solution where small volumes of high-intensity voxels close to the low-intensity region are selected. Such regions may have the lowest average intensities but make the insertion less plausible. For example, the threat may be inserted into an empty corner of a bag but with a small part outside the bag. To address this issue, we have the second item in Eq. (5) which aims to minimize the number of voxels whose values are greater than a threshold c in the selected bag regions. The third item aims to limit the coordinate value in the direction of gravity so that it tends to insert the threat object in a lower location within the bag.

Algorithm 1 Particle Swarm Optimisation for Optimal Insertion

Input: 3D indicator matrix M^{thr} from threat volume segmentation, 3D projection cost map M^{bag} from bag volume segmentation, inertia weight w , cognitive parameter c_1 , social parameter c_2 , number of particles N .

Output: Optimal insertion location and orientation $\Theta^* = \{x^*, y^*, z^*, \alpha^*, \beta^*, \gamma^*\}$.
Set $t = 0$ and randomly initialise Θ_i^0 and V_i^0 for $i = 1, 2, \dots, N$;
while $t \leq T$ **do**
 $t \leftarrow t + 1$;
 Calculate $cost[i]$ using Eq. (5) for $i = 1, 2, \dots, N$;
 Find Θ_i^{best} and Θ^{best} according to the calculated $cost$;
 Update velocities of particles using Eq. (9);
 Update Θ^t using Eq. (8) for each particle;
end while
Output Θ^{best} as Θ^* .

2) *Particle Swarm Optimization:* To solve the problem defined in Eq. (5), one of the enabling methods is particle swarm optimization (PSO) [26], [32]. To make this paper self-contained, we briefly describe the PSO method under our problem setting. The swarm is first initialised with N particles $\{\Theta_i^0\}_{i=1}^N$, where each Θ^i is a vector of six variables. The aim is to find the optimal Θ^* minimizing the objective function defined in Eq. (5) after T iterations. In the t -th iteration, we update the i -th particle Θ_i^t as follows:

$$\Theta_i^t = \Theta_i^{t-1} + V_i^t, i = 1, 2, \dots, N, \quad (8)$$

where

$$V_i^t = wV_i^{t-1} + c_1r_1(\Theta_i^{best} - \Theta_i^{t-1}) + c_2r_2(\Theta^{best} - \Theta_i^{t-1}) \quad (9)$$

is the velocity of i -th particle in the t -th iteration; w, c_1 and c_2 are the inertia weight, cognitive and social parameters respectively [33], [34]; r_1 and r_2 are random numbers drawn from a uniform distribution for each particle in each iteration; Θ_i^{best} and Θ^{best} are the best position of i -th particle and the best position of the swarm (i.e. all particles) thus far respectively.

After T iterations of Eq. (8-9), the optimal variables $\Theta^* = \{x^*, y^*, z^*, \alpha^*, \beta^*, \gamma^*\}$ can be obtained and are ready to use for threat insertion. The algorithm is shown in Algorithm 1.

3) *Image Blending:* Given the optimal insertion location and orientation, a TIP volume can be generated by inserting the segmented threat object volume into the bag volume after rotating it to the optimal orientation. Recall that the segmented threat volume is denoted as I^{thr} , it is firstly weighted by the indicator matrix M^{thr} such that intensity values of voxels in the *uncertain region* are attenuated. We use the same rotation function $f(\cdot)$ in Eq. (3) and the optimal angles derived from PSO to rotate the weighted threat volume:

$$\tilde{I}^{thr} = f(I^{thr} \odot \tilde{M}^{thr} | \alpha^*, \beta^*, \gamma^*). \quad (10)$$

The insertion is an image blending process within 3D imagery in which we modify the values of relevant voxels in the

original bag volume I^{bag} according to the insertion location and rotated volume of the threat. Different methods can be employed for the purpose of image blending. One simple yet effective method is to add the threat volume matrix \tilde{I}^{thr} to the sub-volume matrix of I^{bag} , where the sub-volume is specified by the optimal position x^*, y^* and y^* .

D. TIP Quality Score

Given a threat volume and a baggage volume, our approach is able to generate a TIP volume which may be of variable comparative realism and plausibility given all possible unconstrained combinations of inputs. On one hand, it is well known that particle swarm optimisation could lead to a local best solution [35]. On the other hand, more importantly, a given baggage volume may just not be suitable for a given randomly selected threat to be inserted as the threat is physically too large (e.g. large firearm threat volume into small handbag target volume). As a result, it is important to have a metric evaluating the quality of a TIP generated volume without manual review. Operationally, this can be used to reject poor quality TIP volumes before they are presented to an operator as part of a TIP based performance evaluation system.

We propose the *TIP quality score* to evaluate the quality of generated TIP volumes. Specifically, we use the cost defined in Eq. (5) and normalise it with the volume of inserted threat. The normalised cost value can be easily transformed into a score in the range of 0 – 100 using the following equation:

$$score = \max(0, \min(100, f(cost))) \quad (11)$$

where f is a monotonically decreasing function which could be selected based on the specific requirement of the TIP quality. In our experiment, we use a simple linear function

$$f(cost) = 100 - 0.01 \times cost. \quad (12)$$

E. Metal Artefact Generation

The problem of metal artefacts in X-ray CT images is well studied in medical imaging applications [36], [37], [38]. Metal artefacts are caused by the presence of high-density objects in the scan field of view. The origin of metal artefacts has been studied extensively in the literature and several assumptions have been made [39]. Regardless of the origin of metal artefacts, the effects of these artefacts in the reconstructed CT volumes are the same. Metal artefacts appear as dark and white streaks radiating from the metal objects and spreading across the whole reconstructed CT volumes [13]. They are more prominent near the metal objects and are a function of scan orientation and the material content (see Figure 4).

To enhance the plausibility of generated TIP, it is necessary to take into consideration metal artefacts in the TIP process so that the threat objects appear as if they were genuinely located in the scanned bag. Our proposed metal artefact generation (MAG) procedure depicted in Figure 5 is inspired by the established metal artefact reduction (MAR) projection-replacement techniques in medical imaging applications [37], [38]. In a similar vein to these methods, the whole process of MAG is based on a sequence of 2D slices of a 3D CT volume.

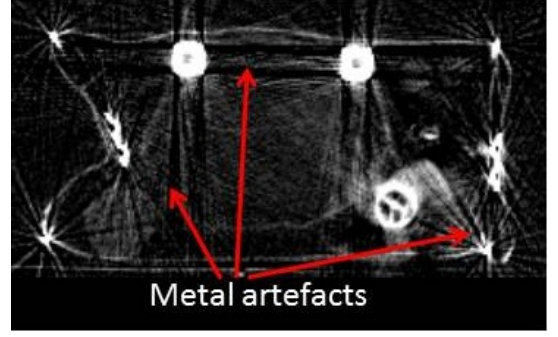


Fig. 4. CT Metal artefacts in a CT slice of a cluttered baggage.

It starts by mapping the original slices of the benign bag (harmless passenger bag), its metal-only slices and the metal-only slices of the artefact-free threat object to the projection domain via the Radon transform [40]. The output of this step is known as a sinogram image.

The artefact free 3D CT volume of a threat object is obtained by appropriate thresholding to remove artefacts and noise. The metal-only volume of a benign bag and the artefact-free threat object are obtained by segmenting the metal objects in their original CT volume by thresholding using a suitable metal CT intensity threshold. This step exploits the fact that metal objects in CT volumes have higher density compared to other objects. Subsequently, the metal traces corresponding to the metal objects of the benign bag and the metal part of the artefact-free threat object are combined in one projection volume. A mask corresponding to all the metal traces is marked in the sinogram of the benign bag CT volume. In conventional MAR projection-replacement based methods, this mask corresponds to the corrupted area in which projection bins are affected by metallic objects and which need to be replaced by surrogate data. Marking this corrupted area in the Radon domain is equivalent to marking all rays passing through the metallic objects originating from the bag and the threat object in the 3D CT TIP volume (benign bag with the threat object). CT metal artefacts emerging from the metallic objects spread across these lines. In order to generate metal artefacts in the benign bag CT volume, the projection bins in the marked mask in the benign bag sinogram are thus made inconsistent with their neighbourhood unlike MAR projection-replacement based methods in which the projection bins in this mask are replaced by interpolated data. The underlying idea behind this is to mimic real CT scanning of a metal object by making the sinogram values corrupted and inconsistent with their neighbourhood if the corresponding X-rays have intersected the metal object. In fact, since metal objects are high-attenuation objects, they heavily attenuate the X-ray beams and consequently, only a few photons reach the scanner detectors. This effect known as photon starvation effect indeed produces corrupted data in the sinogram and gives rise to artefacts in the reconstructed 2D and 3D images. In order to corrupt the projection bins of the marked mask in the benign bag sinogram, we have used an empirical function as follows:

$$s'_{ij} = (1 - q) \times s_{ij} + q \times s_{max}, s_{ij} \in \mathcal{S}_{marked} \quad (13)$$

where s_{ij} and s'_{ij} are the benign bag sinogram values within the marked mask before and after being corrupted, respectively; s_{mask} is the maximum value of benign bag sinogram in the marked mask region; q is a hyper-parameter empirically set to 0.2 in our experiments. As we will show shortly, by following the above steps, consistent metal artefacts are generated within the bag CT volumes which are a function of the scan orientation of the bag, the material of the bag content and the material of the inserted threat object. As depicted in Figure 5, once the metal artefacts are generated in the Radon space, the resulting modified sinogram is re-projected back into the CT domain. The resulting reconstructed CT volume corresponds to the original benign bag CT volume corrupted by metal artefacts originating from the threat object metal part and the benign bag metal objects. The final 3D TIP volume is obtained by combining the resulting CT volume with the artefact free threat object CT volume.

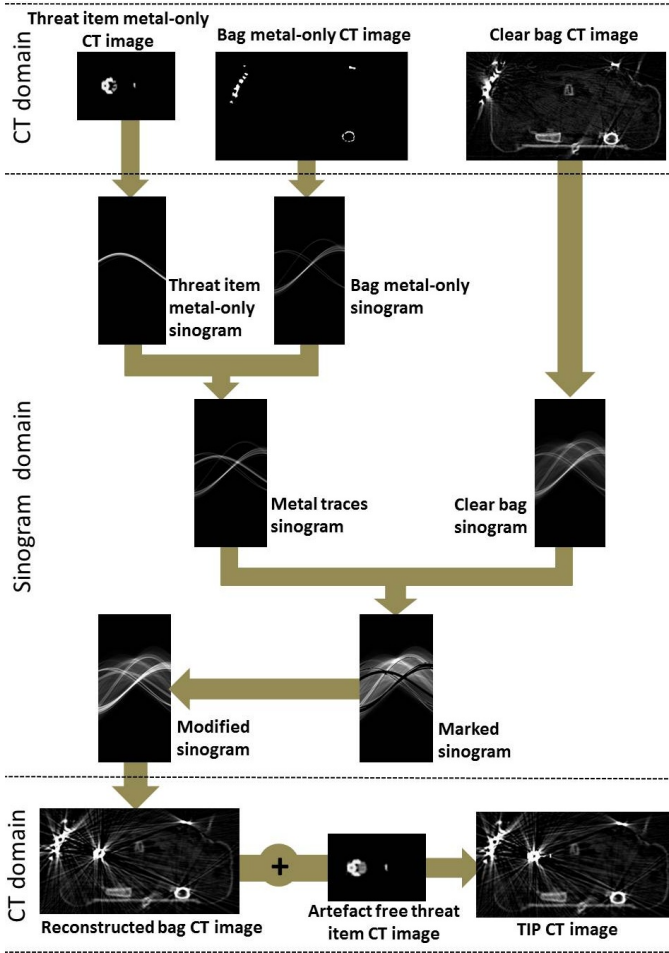


Fig. 5. Flow chart of our MAG method depicted using 2D CT slices

IV. EXPERIMENTS AND RESULTS

In this section, experiments are conducted to evaluate the effectiveness of the proposed approach for 3D TIP in baggage CT imagery. In our experiments, the constant value c in Eq. (2) is set to 100 and the constant value c' in Eq. (6) is set to 10. As a result, voxels with intensity values higher than 410



Fig. 6. Threat volume segmentation results. Rows from top to bottom: bottle, bottle, handgun and submachine gun. Columns from left to right: original volumes, indicator matrix volumes (defined in Eq. (1)) and segmented threat volumes.

(i.e. $10/100 \times 4096$) will be penalised in the particle swarm optimisation. Values of λ_1 and λ_2 in Eq. (5) are empirically set to 0.01 and 1 respectively. For each component of the framework, we present some exemplar results in Figures 6-12 for qualitative performance evaluation. We further generate a large number of TIP using real baggage volumes from an airport for quantitative evaluations.

A. Qualitative Evaluations

Figure 6 shows some exemplary threat volume segmentation results using our proposed algorithm. The original threat volumes of two bottles (in the first two rows) and two firearms (handgun in the third row and submachine gun in the fourth row) are displayed in the left column where the background and noises could be observed. The middle column shows the results of our segmentation algorithm indicated by the 3D matrix defined in Eq. (1), where the threat body regions are coloured in yellow, background regions in white and uncertain regions in grey. The segmented threat volumes are shown in the right column from which we can see the background and noises are removed.

Figure 7 shows bag volume segmentation results using the proposed algorithm. Five different original bag volumes are displayed in the top row and their corresponding segmentation results are shown in the bottom row. The segmentation results are represented by the 3D projection cost maps defined in Eq. (2), where regions of inner-void, bag-content and background

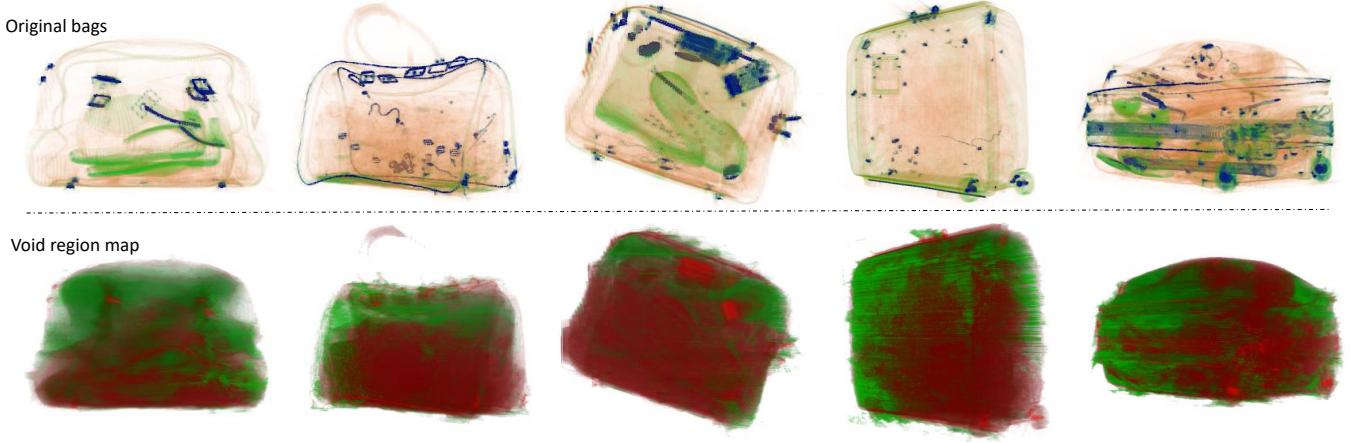


Fig. 7. Baggage volume segmentation results. Top row: original baggage CT volumes; bottom row: projection cost volume defined in Eq. (2) with the green colour representing void regions and the red color representing regions having high projection cost.

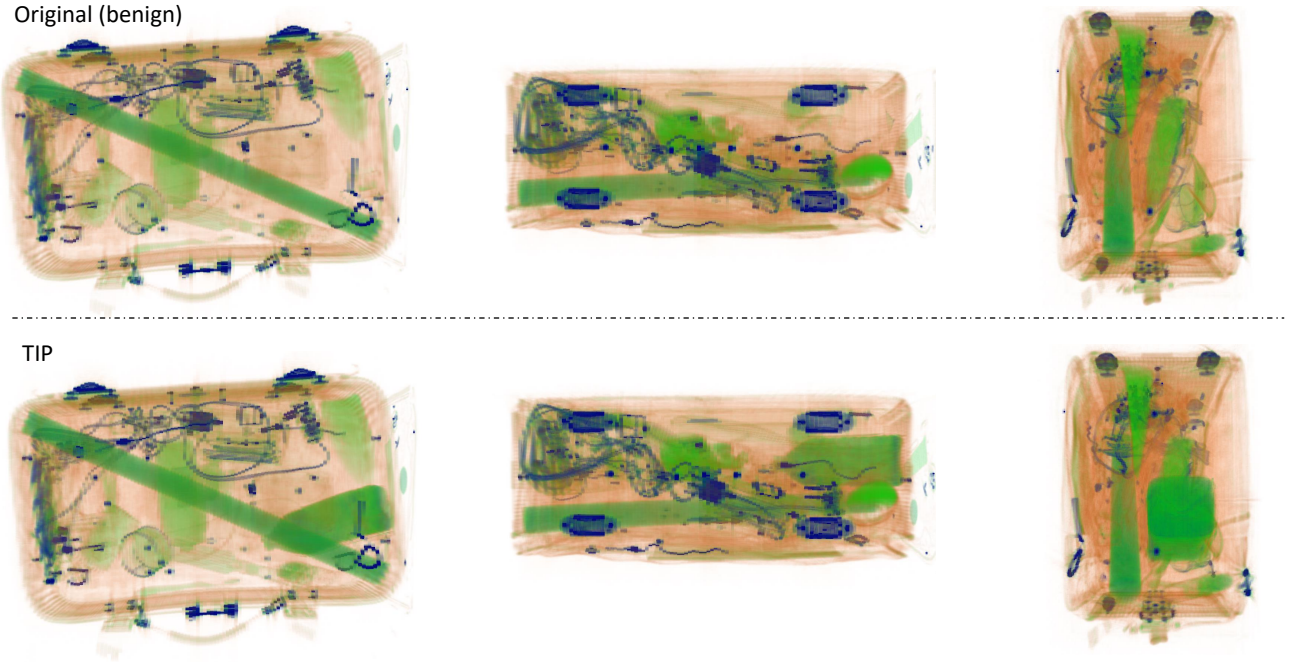


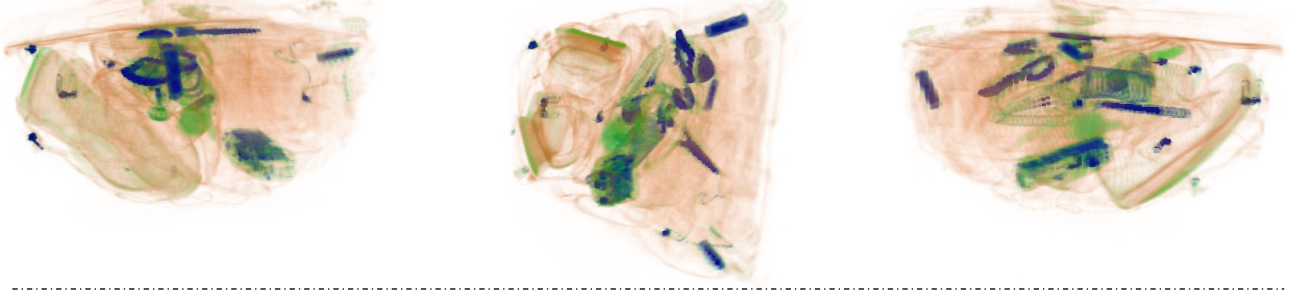
Fig. 8. An exemplar threat image projection (TIP) result where a bottle signature is inserted into a suitcase. Three orthogonal views are shown in three columns. Views of the original baggage and the resultant TIP are shown in the top and bottom rows respectively.

are coloured in green, red and white respectively. The bag segmentation results in Figure 7 indicate our algorithm is able to readily locate accurate bag boundaries as well as the void regions inside the bags.

The results of threat volume projection are shown in Figures 8-11. We visualize the resultant 3D TIP volumes with three orthogonal views in three columns. The first row shows views of the original benign baggage and the TIP results with threat signatures inserted are shown in the second row. We can see that the threat signatures can be successfully projected into the baggage regardless of the volumes and shapes of the threats and baggage. This attributes to a robust segmentation algorithm for threat and baggage segmentation.

Specifically, Figure 8 shows a TIP result with a bottle projected into a cluttered suitcase. Our approach has been successfully discovered the optimal insertion location and orientation and generated a plausible TIP volume. Figure 9 shows a TIP result with a small bottle into a backpack. Although the lack of void region in the original backpack, our algorithm projects the bottle signature into a low-intensity region (orange colour). As a result, the inserted bottle looks like being surrounded by organic materials (e.g., clothes) and very realistic. In figure 10, the signature of a small handgun is inserted to a baggage and Figure 11 shows the TIP result with a submachine gun signature inserted into a very cluttered suitcase. In summary, with satisfying results of threat and bag volume segmentation,

Original (benign)



TIP

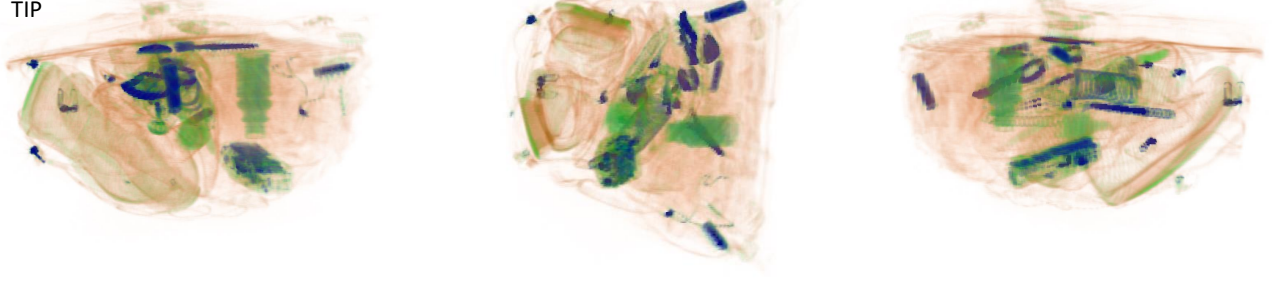
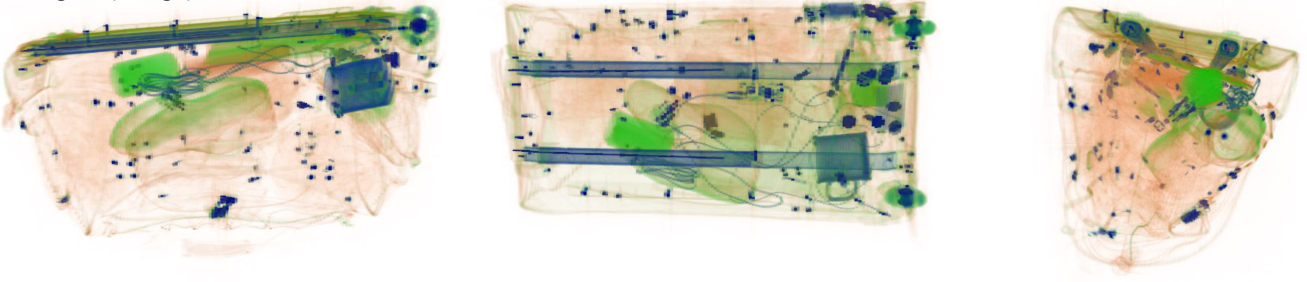


Fig. 9. An exemplar threat image projection (TIP) result where a small bottle signature is inserted into a suitcase. Three orthogonal views are shown in three columns. Views of the original baggage and the resultant TIP are shown in the top and bottom rows respectively.

Original (benign)



TIP

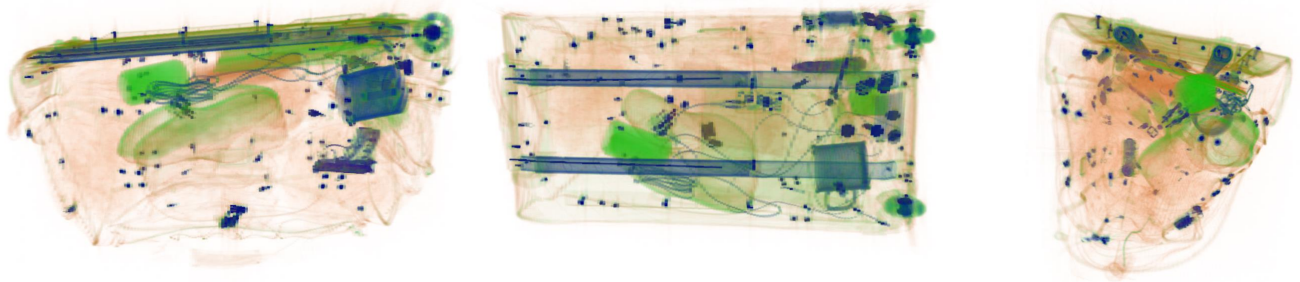


Fig. 10. An exemplar threat image projection (TIP) result where a handgun signature is inserted into a suitcase. Three orthogonal views are shown in three columns. Views of the original baggage and the resultant TIP are shown in the top and bottom rows respectively.

the particle swarm optimisation algorithm is able to find the optimal position and orientation for the insertion hence plausible TIP could be generated as shown in Figures 8-11.

To evaluate the performance of the proposed metal artefact generation algorithm, we present two examples in Figure 12. Selected slices of the original bag volumes are shown in the

first row. Corresponding slices of the TIP without and with MAG are shown in the second and third rows respectively. We can see that slices with MAG in the third row look more realistic in the region of metal objects due to the generation of artefact streaks.

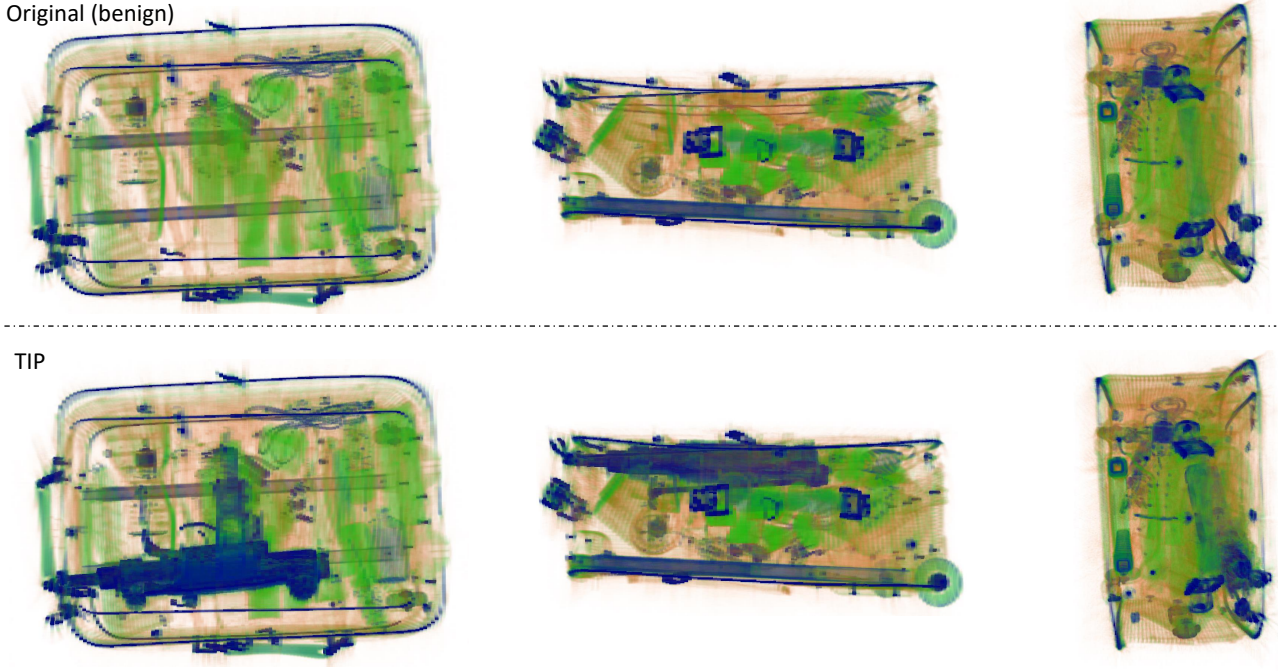


Fig. 11. An exemplar threat image projection (TIP) result where a submachine gun signature is inserted into a suitcase. Three orthogonal views are shown in three columns. Views of the original baggage and the resultant TIP are shown in the top and bottom rows respectively.

B. Quantitative Evaluations

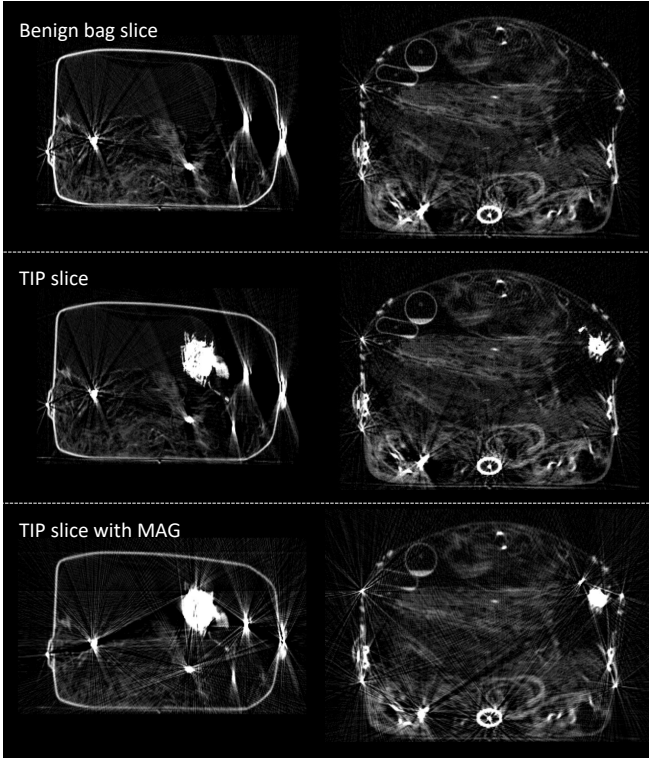


Fig. 12. Exemplar results of metal artefact generation (MAG). Two examples are shown in the left and right columns respectively. Rows from top to bottom: slices of benign bag CT volumes; corresponding slices of TIP volumes without MAG; corresponding slices of TIP volumes with MAG.

We design two experiments to evaluate the proposed TIP approach quantitatively. The first experiment aims to investigate the consistency of TIP quality scores with human evaluations. In Eq. (11), a monotonically decreasing function is used in our experiment. We select 150 generated TIP volumes from a large number of candidates and the scores of selected TIP volumes are evenly distributed in the range of 0 – 100. We use a 3D CT volume visualisation tool to visually inspect each TIP volume and categorize it into one of three classes (i.e. *good*, *medium* and *bad*) according to their quality. A TIP volume is defined as *good* if the threat signature is perfectly inserted into a void region within the bag volume and visually realistic and plausible. A TIP volume is labelled as *bad* if it is obviously unrealistic, for example, the threat signature is inserted outside the bag or intercepted by other items in the bag. A TIP volume of *medium* quality is not perfect but the flaw can only be spotted by careful inspection after the considerable time (> 2 minutes). As a result, there are 102, 37 and 11 TIP volumes labeled as *good*, *medium* and *bad* respectively. It is noteworthy this ratio is not a performance reflection of our TIP approach since we deliberately select low score TIP to look into the relationship between TIP scores and TIP qualities in this experiment. We calculate the mean and standard deviation of the scores for the TIP volumes falling in each class and the results are 82.8 ± 23.2 , 69.8 ± 36.1 and 45.7 ± 48.1 , respectively. In general, the proposed TIP quality scores are consistent with human evaluation results in terms of mean values. On the other hand, however, we also can see large standard deviations of the quality scores for all three classes. It is indicated that the TIP quality score is not

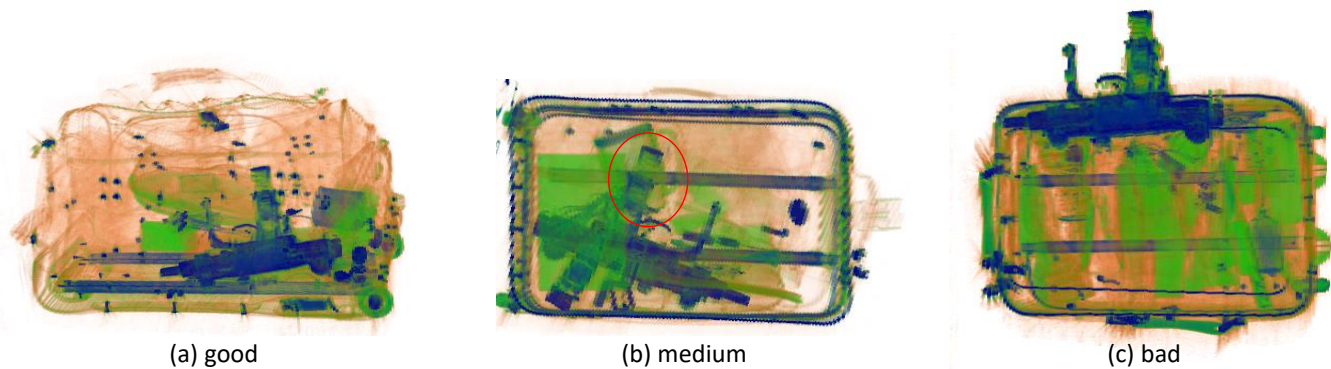


Fig. 13. Examples of *good*, *medium* and *bad* TIP volumes: (a) TIP of good quality; (b) TIP of medium quality due to the interception of the magazine into a mug (highlighted in red circle); (c) TIP of bad quality due to the insertion outside the bag.

perfectly reliable for TIP quality evaluation. As a result, we need to set a high score threshold to reject TIP of bad quality in practice which unavoidably will also falsely reject some good ones.

In the second experiment, we randomly select 100 generated TIP volumes and manually label each of them as the class of *good*, *medium* or *bad* based on visual inspection. As a result, 92% of the TIP volumes are good, 6% are medium and only 2% of them are bad. These results demonstrate that our proposed TIP approach is able to generate TIP volumes of good quality with a very high plausibility and realism acceptance rate.

V. DISCUSSION

In this section, we discuss the limitations of our proposed approach and potential solutions to addressing them in future work. Specifically, we discuss three aspects: preconditions, hyper-parameters and failure cases.

The behaviours of our approach rely heavily on values of many hyper-parameters such as threshold values in volume segmentation and insertion location optimisation. Although the approach is generally robust to most parameters hence performs well as illustrated without exhaustive parameter tuning, there is a subset we need to take special care with when applying this approach to CT volumes captured with different scanners. This is due to the considerable variability of voxel value ranges and noise levels of CT machines from different manufacturers or even different models from the same manufacturer.

One parameter we need to adjust for specific scanners is the threshold value for binarization in the first step of void determination (c.f. Figure 3). This threshold determines the accuracy of bag boundary. We are aware that if this threshold value is too small, the noise surrounding the bag would be mistakenly treated as part of the bag. As a result, a region outside the bag could be potentially treated as inner void and be the place where threats are inserted in. This would make the resultant TIP obviously unrealistic. One way to addressing this issue is to set this threshold value higher, which, however, could mistakenly remove voxels of bag boundaries since the

materials of bag surfaces usually have low intensities in CT volumes. Fortunately, this kind of TIP artefact is more acceptable compared with the former one (i.e. inserting threats outside a bag). We, therefore, suggest a great value rather than a small one for the threshold of binarization in bag volume segmentation.

VI. CONCLUSION

In this paper, we propose an approach to 3D threat image projection for X-ray CT volumes. Qualitative and quantitative evaluations prove that our TIP approach is able to generate realistic and plausible 3D baggage CT volumes containing fictional threat signatures which can be widely used for training baggage screening operators or automatic threat detection models extending of [11]. In our future work, we will investigate how the deep convolutional neural networks can be employed and benefit the performance of bag and threat image segmentation in our TIP approach. On the other hand, it will be more interesting to investigate how the generated TIP volumes could benefit the training of automatic threat detection models as one type of data augmentation strategy [18].

REFERENCES

- [1] S. Akcay, M. E. Kundegorski, C. G. Willcocks, and T. P. Breckon, "Using deep convolutional neural network architectures for object classification and detection within X-ray baggage security imagery," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 9, pp. 2203–2215, 2018.
- [2] S. Akcay, A. Atapour-Abarghouei, and T. Breckon, "Ganomaly: Semi-supervised anomaly detection via adversarial training," in *Proc. Asian Conference on Computer Vision*. Springer, September 2018.
- [3] Q. Wang, K. N. Ismail, and T. P. Breckon, "An approach for adaptive automatic threat recognition within 3D computed tomography images for baggage security screening," *Journal of X-ray Science and Technology*, vol. 28, no. 1, pp. 35–58, 2020.
- [4] "Threat image projection (TIP)," Overseas Territories Aviation Circular, 2017. [Online]. Available: [https://www.airsafety.aero/Requirements-and-Policy/OTACs/OTAR-Part-178-Aviation-Security/Threat-Image-Projection-\(TIP\).aspx](https://www.airsafety.aero/Requirements-and-Policy/OTACs/OTAR-Part-178-Aviation-Security/Threat-Image-Projection-(TIP).aspx)
- [5] "Threat image projection (TIP) library software," Department for Transport, 2016. [Online]. Available: <https://www.gov.uk/government/publications/threat-image-projection-library-software>
- [6] F. Hofer and A. Schwaninger, "Using threat image projection data for assessing individual screener performance," *WIT Transactions on the Built Environment*, vol. 82, 2005.

- [7] V. Cutler and S. Paddock, "Use of threat image projection (TIP) to enhance security performance," in *International Carnahan Conference on Security Technology*. IEEE, 2009, pp. 46–51.
- [8] B. Jansen, "3D scanners can 'digitally unpack' carry-ons and transform airport checkpoints with better, faster security," USA Today, 2017. [Online]. Available: <https://eu.usatoday.com/story/news/2017/11/16/tsa-tests-3-d-scanners-transform-airport-checkpoints-witto-potentially-boost-security-and-speed-line/869982001/>
- [9] N. Hättenschwiler, M. Mendes, and A. Schwaninger, "Detecting bombs in X-ray images of hold baggage: 2D versus 3D imaging," *Human factors*, vol. 61, no. 2, pp. 305–321, 2019.
- [10] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. Van De L., B. Van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Medical image analysis*, vol. 42, pp. 60–88, 2017.
- [11] N. Megherbi, T. Breckon, and G. Flitton, "Investigating existing medical CT segmentation techniques within automated baggage and package inspection," in *Proc. SPIE Optics and Photonics for Counterterrorism, Crime Fighting and Defence*, vol. 8901, no. L. SPIE, October 2013, pp. 1–8.
- [12] N. Megherbi, T. P. Breckon, G. T. Flitton, and A. Mouton, "Fully automatic 3D threat image projection: application to densely cluttered 3D computed tomography baggage images," in *International Conference on Image Processing Theory, Tools and Applications*. IEEE, 2012, pp. 153–159.
- [13] —, "Radon transform based automatic metal artefacts generation for 3D threat image projection," in *Optics and Photonics for Counterterrorism, Crime Fighting and Defence IX; and Optical Materials and Biomaterials in Security and Defence Systems Technology X*, vol. 8901. International Society for Optics and Photonics, 2013, p. 89010B.
- [14] E. Nadler, P. Mengert, and T. Carpenter-Smith, "Airport security screener performance gains due to on-line training and testing," FAA Technical Center, Atlantic City International Airport, NJ, Tech. Rep., 1994.
- [15] J. L. Fobes, S. M. Cormier, D. M. McAnulty, and B. A. Klock, "Operational assessment for screener proficiency evaluation and reporting system (spears) threat image projection." Federal Aviation Administration Washington DC Office Of Aviation Research, Tech. Rep., 1996.
- [16] E. C. Neiderman and J. L. Fobes, "Threat image projection system," May 31 2005, US Patent 6,899,540.
- [17] D. Gudmundson, L. Perron, A. Filiatrault, A. Doyle, and M. Bouchard, "Method and apparatus for providing threat image projection (TIP) in a luggage screening system, and luggage screening system implementing same," Mar. 1 2011, US Patent 7,899,232.
- [18] N. Bhowmik, Q. Wang, Y. F. A. Gaus, M. Szarek, and T. P. Breckon, "The good, the bad and the ugly: Evaluating convolutional neural networks for prohibited item detection using real and synthetically composited X-ray imagery," in *British Machine Vision Conference (BMVC) Workshops*, 2019.
- [19] T. W. Rogers, N. Jaccard, E. D. Protonotarios, J. Ollier, E. J. Morton, and L. D. Griffin, "Threat image projection (TIP) into X-ray images of cargo containers for training humans and machines," in *IEEE International Carnahan Conference on Security Technology*. IEEE, 2016, pp. 1–7.
- [20] D. Neil, N. Thomas, and B. Baker, "Threat image projection in CCTV," in *IEEE International Carnahan Conference on Security Technology*. IEEE, 2007, pp. 272–280.
- [21] F. M. Donald, "CCTV: A challenging case for threat image projection implementation," *Security Journal*, vol. 28, no. 3, pp. 290–308, 2015.
- [22] F. M. Donald and C. Donald, "Vigilance and the implications of using threat image projection (TIP) for CCTVs surveillance operators," in *Australian Security and Intelligence Conference. Perth: Secau*. Citeseer, 2008, pp. 26–36.
- [23] Y. O. Yildiz, D. Q. Abraham, S. Agaian, and K. Panetta, "3D threat image projection," in *Three-Dimensional Image Capture and Applications*, vol. 6805. International Society for Optics and Photonics, 2008, p. 680508.
- [24] Z. Chen, L. Zhang, S. Wang, Y. Sun, Q. Huang, and Z. Tang, "CT system for security check and method thereof," Oct. 3 2018, European Patent EP2960869B1.
- [25] M. Durzinsky, M. A. Morig, and S. König, "Projection of objects in CT X-ray images," 2018, WO 2018/146047 A1.
- [26] J. Kennedy, "Particle swarm optimization," *Encyclopedia of machine learning*, pp. 760–766, 2010.
- [27] A. Rosenfeld, *Digital picture processing*. Academic press, 1976.
- [28] R. Adams and L. Bischof, "Seeded region growing," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 16, no. 6, pp. 641–647, 1994.
- [29] C. R. Maurer, R. Qi, and V. Raghavan, "A linear time algorithm for computing exact euclidean distance transforms of binary images in arbitrary dimensions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 2, pp. 265–270, 2003.
- [30] K. Shoemake, "Animating rotation with quaternion curves," in *ACM SIGGRAPH computer graphics*, vol. 19, no. 3. ACM, 1985, pp. 245–254.
- [31] P. J. Verveer, "An implementation of array rotation in scipy: `scipy.ndimage.interpolation.rotate`."
- [32] J. V. M. Lester, "PySwarms, a research-toolkit for Particle Swarm Optimization in Python," *Journal of Open Source Software*, vol. 3, 2018. [Online]. Available: <https://doi.org/10.21105/joss.00433>
- [33] Y. Shi and R. Eberhart, "A modified particle swarm optimizer," in *IEEE International Conference on Evolutionary Computation*. IEEE, 1998, pp. 69–73.
- [34] J. C. Bansal, P. K. Singh, M. Saraswat, A. Verma, S. S. Jadon, and A. Abraham, "Inertia weight strategies in particle swarm optimization," in *World Congress on Nature and Biologically Inspired Computing*. IEEE, 2011, pp. 633–640.
- [35] P. J. Angeline, "Evolutionary optimization versus particle swarm optimization: Philosophy and performance differences," in *International Conference on Evolutionary Programming*. Springer, 1998, pp. 601–610.
- [36] B. D. Man, J. Nuyts, P. Dupont, G. Marchal, and P. Suetens, "Metal streak artifacts in X-ray computed tomography: a simulation study," *IEEE Transactions on Nuclear Science*, vol. 46, no. 3, pp. 691 – 696, 1999.
- [37] A. Mehranian, M. Ay, A. Rahmim, and H. Zaidi, "Sparsity constrained sinogram inpainting for metal artifact reduction in X-ray computed tomography," *IEEE Nuclear Science Symposium and Medical Imaging Conference, Barcelona, Spain*, pp. 3694–3699, 2011.
- [38] —, "Metal artifact reduction in CT-based attenuation correction of PET using sobolev sinogram restoration," *IEEE Nuclear Science Symposium and Medical Imaging Conference, Barcelona, Spain*, pp. 2936–2942, 2011.
- [39] A. Mouton, N. Megherbi, K. Van Slambrouck, J. Nuyts, and T. P. Breckon, "An experimental survey of metal artefact reduction in computed tomography," *Journal of X-ray Science and Technology*, vol. 21, no. 2, pp. 193–226, 2013.
- [40] S. R. Deans, "The radon transform and some of its applications," *New York: John Wiley & Sons*, 1983.